Adaptation and Speciation in Genetic Modeling of Physical Systems

15 September 2009

Albert H. Zimmerman
Electronics and Photonics Laboratory
Physical Sciences Laboratories

Prepared for:

Space and Missile Systems Center
Air Force Space Command
483 N. Aviation Blvd.
El Segundo, CA 90245-2808

Authorized by: Engineering and Technology Group

AEROSPACE
Assuring Space Mission Success

David E. Davis
SMC/EA

REPORT DOCUMENTATION PAGE

1. REPORT DATE (DD-MM-YYYY)	2. REPORT TYPE	3. DATES COVERED (From - To)
15-09-2009		

4. TITLE AND SUBTITLE	5a. CONTRACT NUMBER
Adaptation and Speciation in Genetic Modeling of Physical Systems	FA8802-09-C-0001
	5b. GRANT NUMBER
	5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)	5d. PROJECT NUMBER
Albert H. Zimmerman	**5e. TASK NUMBER**
	5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)	8. PERFORMING ORGANIZATION REPORT NUMBER
The Aerospace Corporation Physical Sciences Laboratories El Segundo, CA 90245-4691	TR-2009(8550)-11

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)	10. SPONSOR/MONITOR'S ACRONYM(S)
Space and Missile Systems Center Air Force Space Command	SMC
483 N. Aviation Blvd. El Segundo, CA 90245	**11. SPONSOR/MONITOR'S REPORT NUMBER(S)**

12. DISTRIBUTION/AVAILABILITY STATEMENT

Approved for public release; distribution unlimited.

13. SUPPLEMENTARY NOTES

20100302249

14. ABSTRACT

A genetic algorithm has been used to allow a population of first principles lithium-ion (Li-ion) battery cell models to adapt to a database obtained from life-testing 50-Ah Li-ion cells. This adaptive process allows the fundamental processes that control the cell performance to be extracted from the life test data. During the process of genetic adaptation, the population of battery cell models has been observed to separate into two distinct species of models. The emergence of separate species is recognized by the inability to produce viable offspring models from mixtures of the genetic codes for the two species. In this situation, the two species are found to co-exist for up to many hundreds of generations. However, one of the two species eventually suffered extinction because it did not retain adequate genetic diversity and because its fitness did not continue to improve as did that of the other species. The dynamics involved in genetic adaptation of battery cell models, including the roles of speciation and extinction, are described.

15. SUBJECT TERMS

Battery, Lithium ion, Model, Genetic analysis, Speciation, Life test, Adaptive modeling, Life prediction

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	**b. ABSTRACT**	**c. THIS PAGE**	Leave blank	15	Albert Zimmerman
UNCLASSIFIED	UNCLASSIFIED	UNCLASSIFIED			**19b. TELEPHONE NUMBER** *(include area code)* (310)336-7415

Contents

Figures

1. Introduction

Genetic modeling of physical systems provides a technique that enables complex models of these systems to undergo adaptation such that they realistically describe the behavior of the physical system. This modeling method utilizes a genetic algorithm to adapt a population of models to the observable characteristics of the physical system. The fittest models in the population are those that best describe the physical system, and the genetic algorithm enables the fittest models to most strongly influence the emergence of each generation in the population from the previous generation. This technique enables complex models that contain dozens of adjustable parameters to find the combination of parameters that most accurately describes the real-world physical system that is being modeled.

Here we use a first principles model of a lithium-ion (Li-ion) battery cell with a genetic algorithm to determine the set of parameters that best matches the observed performance of a 50-Ah battery cell during a life test. The dynamics of how the population of models adapts to the data will be examined. The roles of the initial population distribution and the population diversity during the genetic adaptation process will be discussed. The discussion will highlight the conditions that can lead to the spontaneous formation of multiple species of models, as well as the conditions that can lead to extinction of species within the population.

2. Model Description

The model used here to describe the physics and chemistry of a prismatic lithium-ion (Li-ion) battery cell and predict its performance from first principles has been described in detail in Reference 1. This model couples a finite-element description of the macroscopic cell geometry to a first-principles description of the key processes that control cell performance. Nine key parameters in the model that are expected to potentially change during a life test are used to comprise the genome that is subject to genetic adaptation. These nine parameters are:

1. Total cathode capacity for Li ions

2. Initial cell state of charge

3. Lithium-ion diffusion rate in the cathode active material

4. Total anode capacity to reversibly intercalate lithium

5. Amount of charge remaining in anode or cathode when the cell is fully discharged (anode/cathode charge balance)

6. Charge transfer resistance in cathode

7. Cell series resistances (anode SEI and electrolyte resistance)

8. Cathode surface polarization rate

9. Cathode surface polarization potential

3. Genetic Algorithm

Each of the nine parameters described in the previous section constitutes one of the nine genetic codons that describe an individual model of the lithium-ion (Li-ion) cell. The genetic analysis performed here involved a population of 40 individual models, each having its own individual genome. The starting population utilized the 40 models that best fit the data from a group of 400 models that were randomly generated using physically reasonable ranges for each of the codons. The fitness of each specific model was based on the average fit between the cell voltage data during a charge/discharge cycle and that predicted by the model based on its genome. A genetic algorithm was used to track the lifetime of each model in the population, and defined the offspring produced by each model to form the next generation, as well as the epigenetic shifts in the genome that could occur during the lifetime of each individual model.

The genetic algorithm has been defined in detail in Reference 1. The key elements that this algorithm uses to define each generation of models from the previous generation are:

1. Epigenetic drift was allowed to occur over the lifetime of an individual model to improve the fitness of its genome in response to its local environment.

2. Roulette selection (selection probability proportional to fitness) provided the basis for selecting which model was to generate an offspring model, and for selecting a model with which it could mate.

3. Genetic mixing was used to produce offspring, which was based on selection of a random point along the 9-dimensional linear vector connecting the genomes of two individual models and extending 50% past each individual genome.

4. Mutations occurred at the birth of 0.5% of the offspring, which involved random shifts in a randomly selected codon within the physically reasonable range for that codon. This helped maintain genetic diversity in the population.

5. Only offspring having fitness better than their parents were included in the models that made up each new generation. All other offspring were allowed to die without contributing to propagation of the next generation.

6. If an offspring model having improved fitness could not be generated after 80 mating attempts, the parent model was preserved in the new population. This rarely occurred.

7. Each new generation was limited to 40 offspring.

As the genetic algorithm described above changed the population of models from generation to generation, the quality of the fit between the cell test data and the model prediction of the cell voltage improved. The adaptation process was assumed to have converged to the optimum model when the deviation of the fittest model in the population improved by 0.000002 V or less over four full generations. For the life test data analyzed in this report, the average deviation between the data points and

the model performance prediction at convergence was in the range of 0.004 to 0.008 V. A typical example showing the model results compared to the data at convergence is shown in Figure 1.
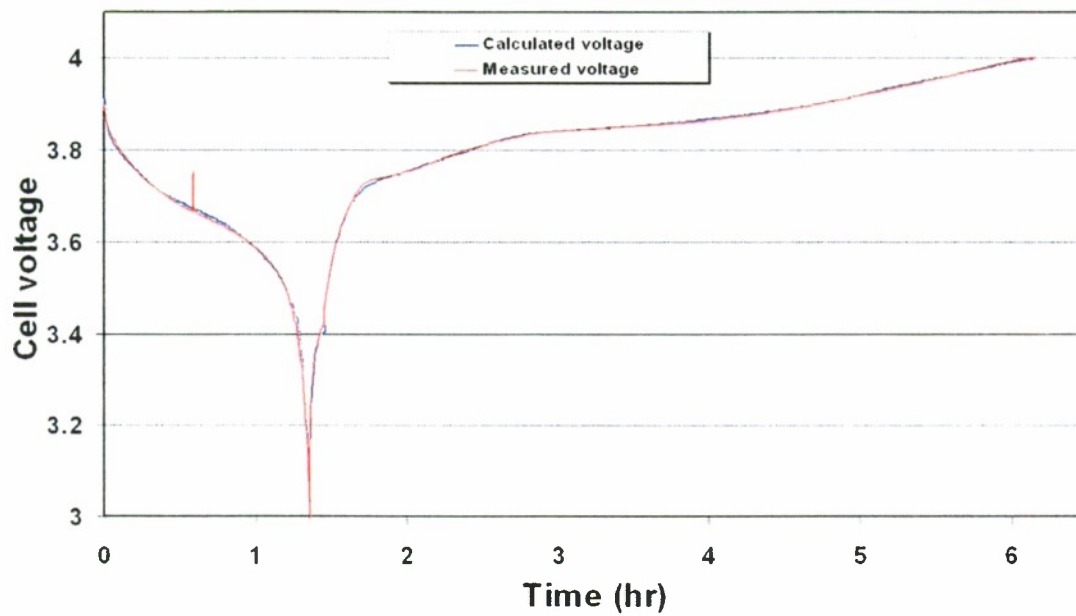

Figure 1.   Measured and optimized model cell voltages after 5505 cycles at 40% depth of discharge.

4. Typical Genetic Adaptation Dynamics

The process of genetic adaptation can be followed by examining the fitness distribution of the population of models over the generations as adaptation continues to improve the quality of the fit to the data. This shift in distribution is shown in Figure 2 for a typical process of genetic adaptation to a test dataset. Each distribution has a characteristic width that reflects the genetic diversity of the population. The fittest model in each population occurs at the lower threshold of the distribution, which is typically somewhat below the most probable fitness in the population.

A number of features in Figure 2 are of interest. First, the starting distribution tends to be relatively wide since the starting models are randomly generated over a range of starting parameters that is wide enough to cover all likely parameter values. As shown by the blue curve in Figure 2, after 40 generations, the population has a fitness of 15 to about 30 mV. As the fitness of the population improves over the generations, the most noteworthy feature is the significant narrowing of the distribution, until eventually the distribution is extremely narrow after convergence to the optimum model. For the case of Figure 2, it has been verified that convergence occurred at the true optimum fit by running the analysis several times, each time finding convergence to the same genetic code. It is important to the adaptation process that the distribution width does not become too narrow until the optimum solution is approached; otherwise, the genetic analysis could be trapped into a false solution. The maintenance of adequate genetic diversity is the key to assuring that the parameter distributions remain wide enough to explore all possible values for each parameter in the genome.
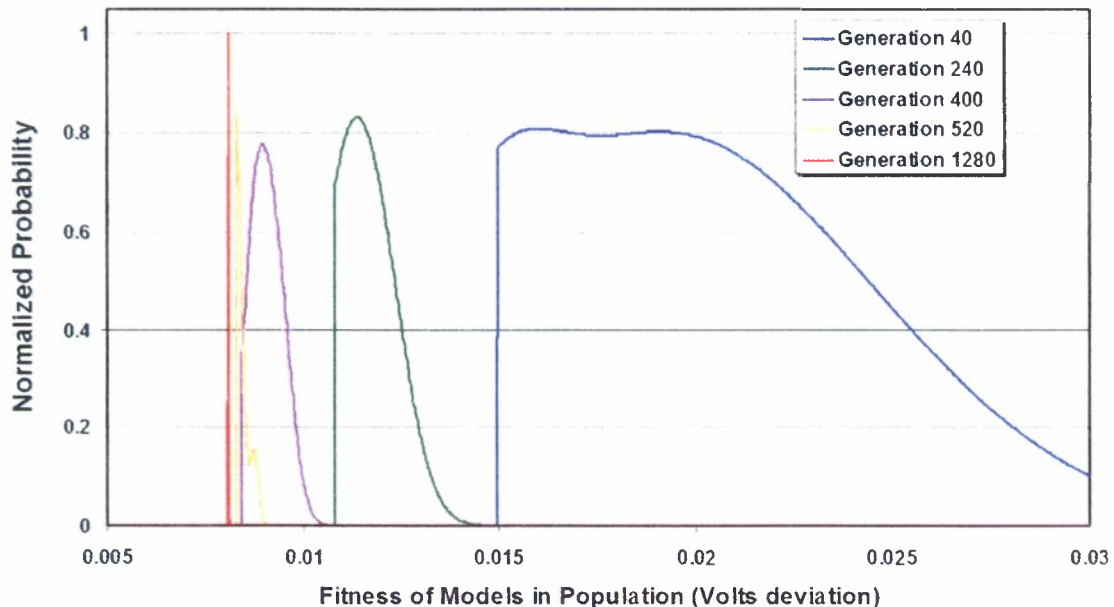

Figure 2.   Change in the fitness distribution through the generations of a genetic algorithm for a typical dataset following 7538 cycles of a 50-Ah Li-ion cell at 30% depth of discharge. The sharp red distribution represents the population after convergence to the optimum solution.

Another feature of interest in Figure 2 is the occasional suggestion of a bimodal distribution, seen in the blue curve and in the gold curve. These features are caused by the random concentration of population members in specific ranges of parameters as new generations are born. Normally these features persist for only a few generations before being overwhelmed by the generational population shifts.

In the case of Figure 2, convergence was relatively rapid because the converged values for each of the nine genetic codons were well within the initial range specified for each of these parameters. If the converged value for a particular codon is well outside the initial specified range, a somewhat different approach to convergence is found, as is shown in Figure 3. In Figure 3, the fittest model in the population is shown as a function of generation. In this case, we see a rapid improvement in fitness down to about 0.004 V over the first 300 generations, followed by a very slow convergence over about 3000 generations to the final solution. The reason for the slow convergence is because the correct cathode capacity was about 10 Ah higher than the range for cathode capacity that was initially specified. During the 3000 cycles of slow convergence, the cathode capacity gradually walked upward about 10 Ah to the correct value. This example illustrates not only how important the initial parameter ranges are, but also the power of the genetic algorithm in finding an optimized value that was well outside the initially expected range.
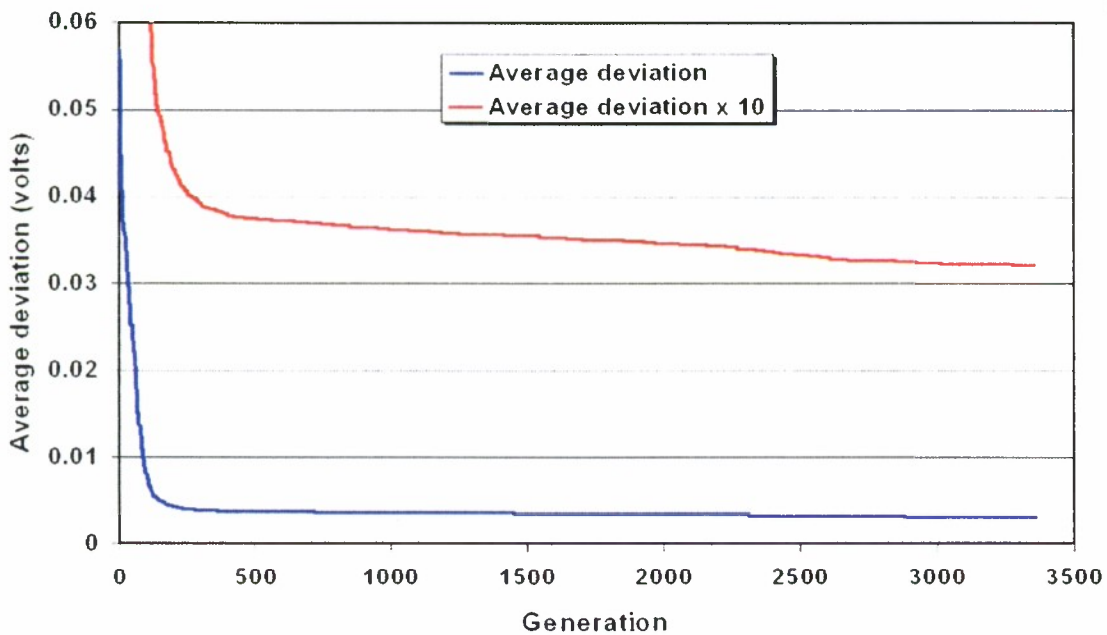

Figure 3.   Approach of the fittest model in the population distribution toward the optimum as a result of genetic adaptation. The long convergence tail results because the allowed cathode capacity range was initially chosen to be 10 Ah below the correct cathode capacity.

5. Speciation during Genetic Adaptation

The process of genetic adaptation does not always proceed as smoothly as illustrated in Figures 2 or 3. Figure 4 illustrates the changes in the distributions of population fitness obtained from a genetic analysis during which a bimodal population distribution appeared and then persisted for hundreds of generations. The bimodal distribution appeared to result from the formation of a new species of model that was distinct from the main population. The new species of model, which is shaded in blue in Figure 4, was characterized by a complete inability to mate with models in the original species and produce viable offspring. All offspring of such mating encounters were so extremely unfit that they had no hope of surviving. Thus, the main species adapted by mating with models of its own species, and the new species adapted only through mating with others of its own species. When the new species originally formed, it contained up to 18 members of the entire population of 40.

The new species continued to thrive for hundreds of generations after its initial formation several hundred generations into the analysis. However, by generation 720 it was beginning to show signs of genetic weakness. The number of the new species individuals had dropped down to about 12 in each generation, and its genetic diversity had begun to become seriously reduced. By generation 920, essentially all genetic diversity had been lost, and only 9–10 individuals remained. The last 5 individuals in this new species failed to produce any viable offspring after generation 1040, and the new species was thereafter extinct.
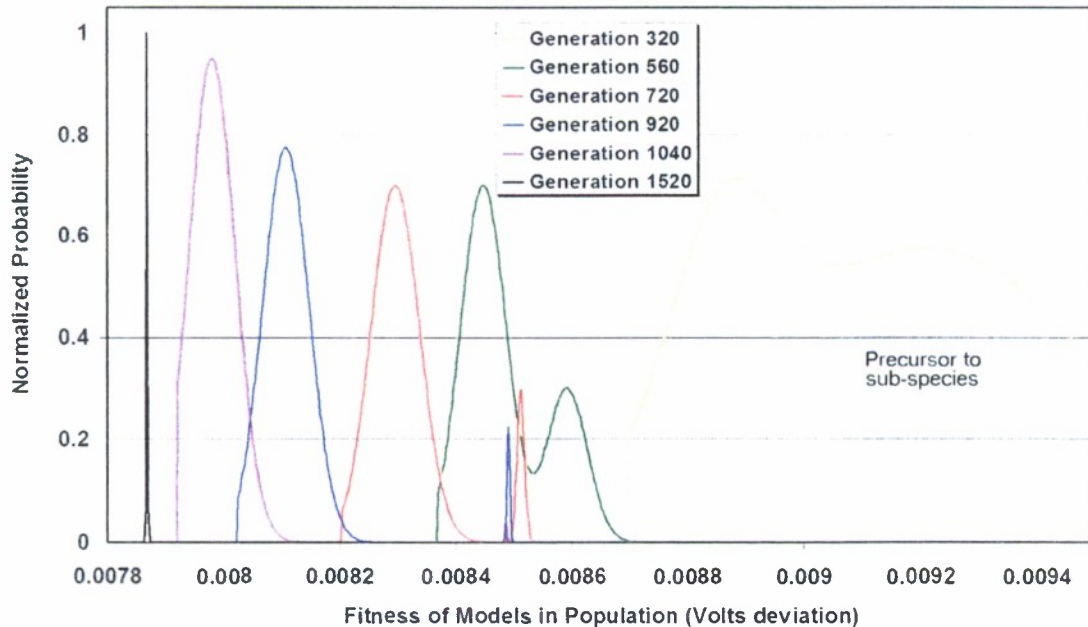

Figure 4. Evolution of a bimodal fitness distribution during the genetic analysis of a dataset after 3474 cycles of a 50-Ah Li-ion cell at 30% depth of discharge.

5.1   Causes for Speciation

A new species seems to be characterized by several key properties. The first of these is that its genome is significantly different from that of other species. The second key characteristic is that its members can only produce viable offspring by mating with member of their own species. The formation of a new species, as well as its later slide into extinction, is of general interest in terms of the underlying causes for these events. The formation of the new species could in principle be the result of either some unusual characteristic of the dataset that was being analyzed, or the result of some unusual asymmetry in the starting model distribution that became accentuated during the genetic analysis.

The dataset that was being analyzed when the speciation was observed was compared to other datasets from the same cell at different points in time during the life test. The datasets appeared very similar, and no features could be found that would be expected to produce different behavior for one and not the others. A repeated genetic analysis of the dataset that caused the speciation in Figure 4 gave no evidence of speciation. Thus, the conclusion was reached that the underlying cell dataset was not the cause for the speciation.

The other likely cause was an unusual distribution of parameters in the starting population selected for the genetic analysis. The starting population contained the 40 fittest models from 400 randomly generated models. The distribution of the fitness of the starting models for the analysis in which the new species was seen is shown in Figure 5, along with the starting population for two other similar analyses.

The results in Figure 5 do suggest that there was at least one, and perhaps two sub-distributions that were present in the initial population for capacity check 3, and that later developed into distinct species of models. To determine the root cause for this separation into species, we examined the individual parameter distributions that developed as the new species formed. While a number of the nine parameters that make up the genome for the model developed bimodal distributions as the new species formed, the key parameter that appeared to prevent viable mating between the species was the SEI conductivity in the anode of the Li-ion cell.
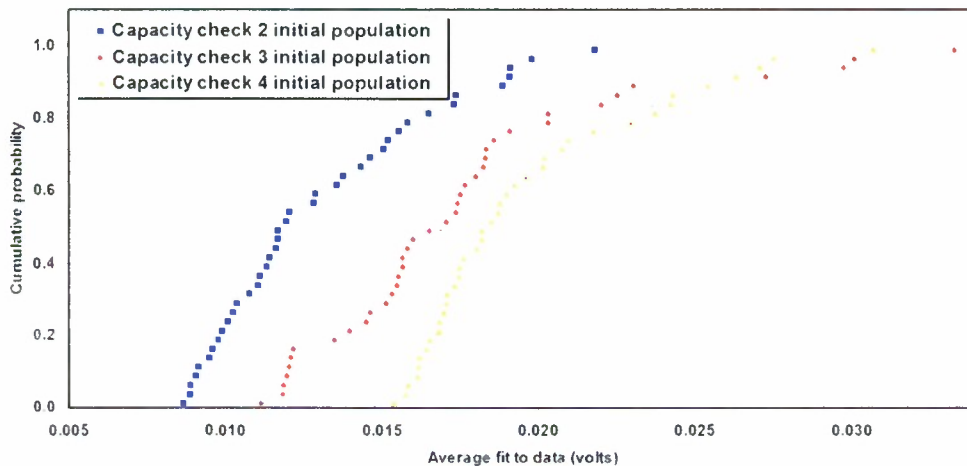

Figure 5.   Initial fitness distributions for starting 40 models in genetic analysis. Capacity check 3 was the analysis for which the population segregated into two distinct species.

As shown in Figure 6, after about 100 generations, the models had separated into one group having low SEI conductivity and another group having high SEI conductivity. The high-conductivity group was the normal species that continued to adapt and eventually converged to the optimum solution. The low-conductivity group became the new species that survived for hundreds of generations before becoming extinct. Careful examination of the genetic algorithm showed why these species went down their separate evolutionary paths. Because the two species developed more than an order of magnitude difference in SEI conductivity, the vast majority of the genetic mixtures that resulted from mating gave an SEI conductivity that was well below the low distribution range, well above the high distribution range, or somewhere in between. None of these regions gave a model fitness that was competitive with the existing models in either distribution, thus causing essentially all inter-species offspring to die without reaching maturity. Clearly, this situation met both criteria for having two distinct species: both significantly differing genomes, and loss of the ability to mate with the other species.

5.2   Causes for Extinction

Once a stable species has co-existed within a population of models for hundreds of generations, the reasons for the extinction of that species are of interest in terms of understanding the dynamics of genetic algorithm operation. For the two species that appeared in the example described above, three factors appear to conspire against the survival of the new species that formed and then eventually disappeared:

1. The species lost the genetic diversity needed to enable it to adapt to the changing popula-
    tion environment.

2. Its competing species became significantly more fit over time, providing a reproductive
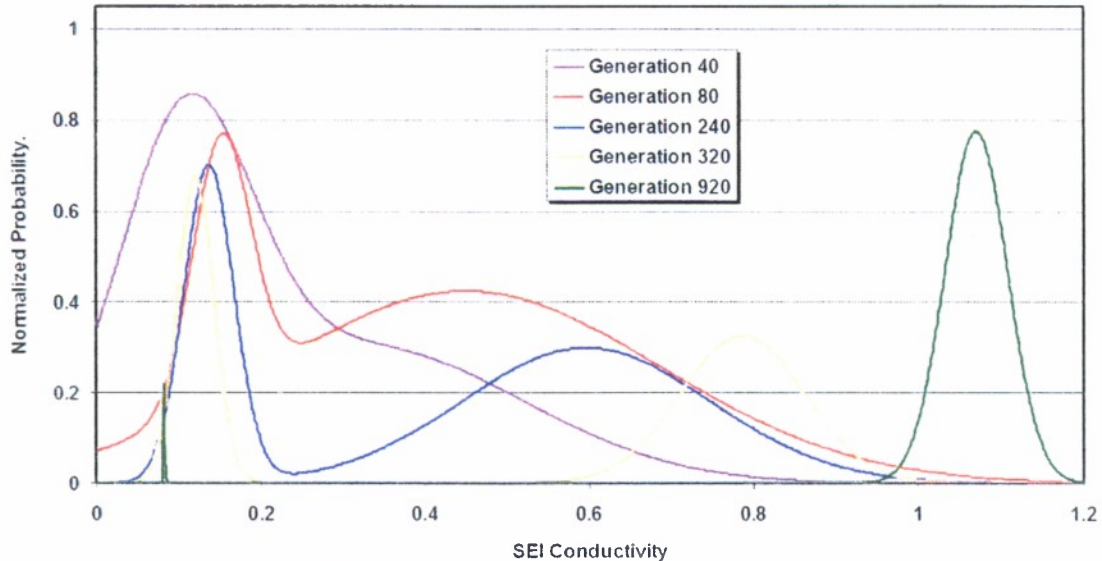   advantage.

Figure 6.   Distributions of SEI conductivity observed during the genetic analysis of the data from capacity check 3 after 3474 cycles of a 50-Ah Li-ion cell at 30% depth of discharge.

11

3. The number of individuals fell below a threshold needed to assure success in finding a mate and producing viable offspring for each generation.

The first of these factors, loss of genetic diversity, is recognized by the extremely narrow distribution for the low-conductivity species in Figure 6 at 920 generations, compared to the much wider distribution for the high-conductivity species. The same shifts in diversity can also be seen in the fitness distributions of Figure 4. This loss of diversity is a signal that this species has reached a genetic dead end, from which it cannot continue to significantly improve its fitness to compete with other species.

The second factor, suffering from a fitness disadvantage, is partially a consequence of the loss of the ability to continue adapting. This, combined with the ever improving fitness of the species that is competing for reproductive success, makes mating opportunities ever more sparse for the less fit species (selection for mating and number of offspring are in direct proportion to fitness). Thus, it is the continued improvement in the fitness of the one species that eventually puts the other species at a significant reproductive disadvantage. The extinction probably would never have occurred if both species had remained with about the same fitness.

The final factor, inability to sustain a critical number of species members, stems from the reproductive disadvantage described above. Since the species can only mate successfully with other members of the same species, the probability that at least two viable offspring will be produced each generation decreases significantly below a population size threshold. For the population size and conditions here, the extinction occurred after about 100 generations with a population size between 8 and 9 individuals. Any species that produces less than two offspring in a generation will become extinct in the next generation, for the genetic algorithm used here.

6. Conclusions

The dynamics of adaptation using a genetic algorithm point to several conclusions that are important to understand how genetic optimization of physical systems can be done most reliably and efficiently. First, the results show how robust the genetic algorithm is for finding the correct solution, even when incorrect initial parameter ranges or bimodal input parameter distributions are provided. While specifying an initial cathode capacity that was 10 Ah off made convergence to the correct capacity very slow, the correct capacity was eventually obtained. While specifying a quite wide range for SEI conductivity contributed to a bimodal species distribution, the more fit species eventually forced the extinction of the less fit one, and the correct conductivity was eventually obtained.

The final conclusion from this study has to do with the importance of the initial ranges for the parameters that comprise the genome for the model. Wide ranges for these parameters are desirable because we do not always know the parameters closely. However, wide ranges can significantly increase convergence time and can increase the probability of converging to an incorrect solution. Narrow initial parameter ranges increase the probability of either missing the correct solution because it is outside the range that is available to genetic adaptation, or substantially increasing the computation time needed to converge to the correct solution.

The best approach for choosing initial parameter ranges is probably one that mimics how an actual life test is run. Start at the beginning of the life test with relatively wide ranges for the model parameters. For each incremental point in the life of a cell, use the previous optimum parameters as the starting point (with a narrower range around them) for analysis of the next point in the lifetime of the cell. This approach is expected to minimize the computation time involved in the genetic analysis of the life test data from a battery cell.

13

Reference

1. Albert. H. Zimmerman, *Genetic Modeling of Battery Performance*, The Aerospace Corporation Technical Report, in press.

PHYSICAL SCIENCES LABORATORIES

The Aerospace Corporation functions as an "architect-engineer" for national security programs, specializing in advanced military space systems. The Corporation's Physical Sciences Laboratories support the effective and timely development and operation of national security systems through scientific research and the application of advanced technology. Vital to the success of the Corporation is the technical staff's wide-ranging expertise and its ability to stay abreast of new technological developments and program support issues associated with rapidly evolving space systems. Contributing capabilities are provided by these individual organizations:

Electronics and Photonics Laboratory: Microelectronics, VLSI reliability, failure analysis, solid-state device physics, compound semiconductors, radiation effects, infrared and CCD detector devices, data storage and display technologies; lasers and electro-optics, solid-state laser design, micro-optics, optical communications, and fiber-optic sensors; atomic frequency standards, applied laser spectroscopy, laser chemistry, atmospheric propagation and beam control, LIDAR/LADAR remote sensing; solar cell and array testing and evaluation, battery electrochemistry, battery testing and evaluation.

Space Materials Laboratory: Evaluation and characterizations of new materials and processing techniques: metals, alloys, ceramics, polymers, thin films, and composites; development of advanced deposition processes; nondestructive evaluation, component failure analysis and reliability; structural mechanics, fracture mechanics, and stress corrosion; analysis and evaluation of materials at cryogenic and elevated temperatures; launch vehicle fluid mechanics, heat transfer and flight dynamics; aerothermodynamics; chemical and electric propulsion; environmental chemistry; combustion processes; space environment effects on materials, hardening and vulnerability assessment; contamination, thermal and structural control; lubrication and surface phenomena. Microelectromechanical systems (MEMS) for space applications; laser micromachining; laser-surface physical and chemical interactions; micropropulsion; micro- and nanosatellite mission analysis; intelligent microinstruments for monitoring space and launch system environments.

Space Science Applications Laboratory: Magnetospheric, auroral and cosmic-ray physics, wave-particle interactions, magnetospheric plasma waves; atmospheric and ionospheric physics, density and composition of the upper atmosphere, remote sensing using atmospheric radiation; solar physics, infrared astronomy, infrared signature analysis; infrared surveillance, imaging and remote sensing; multispectral and hyperspectral sensor development; data analysis and algorithm development; applications of multispectral and hyperspectral imagery to defense, civil space, commercial, and environmental missions; effects of solar activity, magnetic storms and nuclear explosions on the Earth's atmosphere, ionosphere and magnetosphere; effects of electromagnetic and particulate radiations on space systems; space instrumentation, design, fabrication and test; environmental chemistry, trace detection; atmospheric chemical reactions, atmospheric optics, light scattering, state-specific chemical reactions, and radiative signatures of missile plumes.